The Trade Off Between Efficiency, Effectiveness and Social Feasibility of Regulating Road Transport Externalities

Verhoef, E.T.; Nijkamp, P.; Rietveld, P.

published in
Transportation Planning & Technology
1996

DOI (link to publisher)
10.1080/03081069608717572

Transportation Planning and Technology, 1996, Vol. 19, pp. 247–263
Reprints available directly from the Publisher
Photocopying permitted by license only

THE TRADE-OFF BETWEEN EFFICIENCY, EFFECTIVENESS, AND SOCIAL FEASIBILITY OF REGULATING ROAD TRANSPORT EXTERNALITIES

ERIK VERHOEF, PETER NIJKAMP and PIET RIETVELD

Department of Spatial Economics, Free University, Amsterdam, The Netherlands

(Received August 17, 1995; in final form January 19, 1996)

This paper discusses the tension between the theory and practice of the regulation of road transport externalities from the viewpoint of the trade-off between efficiency, effectiveness and the social feasibility of regulation. Various possible types of regulatory instruments, subdivided into 'direct' demand management, 'indirect' demand management and 'supply side' oriented policies, are evaluated qualitatively according to these criteria. Attention is then directed towards the trade-off between efficiency, effectiveness and feasibility within the field of 'direct' demand regulation of road transport.

KEY WORDS: Regulation, externalities, road transport, efficiency, feasibility.

1. INTRODUCTION

Transport policy offers a glaring example of a field with a large discrepancy between theory and practice. One of the main instances of this gap is the fact that efficient pricing instruments for the optimal regulation of road transport externalities, such as electronic road pricing (ERP), apparently provoke so much social and political resistance that they are not likely to be widely introduced and accepted, at least not in the short run. Recent publications on road pricing have consequently focused on the question of its social and political feasibility.[1,2,3,4,5] Others have concentrated on methods of introducing road pricing schemes.[6,7,8,9,10] Finally, much interest has also arisen in various alternatives to ERP in traffic demand regulation.[11,12] Examples are fuel taxes,[13] parking policies,[14,15,16] and tolling with untolled alternatives.[17]

In this paper we will discuss the tension between the theory and practice of the regulation of road transport from the viewpoint of the trade-off between efficiency, effectiveness and the social feasibility of regulation. Besides congestion, which is usually central in such analyses, we will also consider other external costs of road transport (noise, pollution, accidents). Furthermore, we will broaden the scope from 'direct' demand management towards a framework in which also other regulatory instruments ('indirect' demand management and 'supply side' oriented policies) càn be evaluated.

The paper is structured as follows. Section 2 addresses the fundamental trade-off between efficiency and the feasibility of regulation in a simple model. In Section 3, the efficiency, effectiveness and feasibility of the various instruments for containing road transport externalities are discussed. Next, in Section 4, we focus on the trade-off between efficiency, effectiveness and the social feasibility of what we call 'direct demand management'. Finally, Section 5 contains some concluding remarks.

2. THE FEASIBILITY OF EFFICIENCY AND THE EFFICIENCY OF FEASIBILITY IN REGULATION: A BASIC ANALYSIS

Road transport practically always causes external costs. These can be subdivided into *intra-sectoral externalities*, which road users pose upon one another (notably congestion and, to some extent, accidents), and *environmental externalities*, which are shifted to society at large (pollution, noise and also accidents).[18] Because of these external costs, there is overconsumption of road transport: the free market outcome is not equal to the Pareto optimal level of road mobility. This is illustrated in Figure 1. Along the horizontal axis, mobility (for instance, measure in vehicle-kilometres) is depicted; the vertical axis measures costs and benefits in monetary values. The market equilibrium N^0 is at the intersection of the demand curve, which is equal to the marginal private and social benefits (D=MPB=MSB),[19] and the marginal private cost curve (MPC). With identical road users, MPC may be equated to average social cost (ASC); it is positively sloped because of intra-sectoral externalities.

Taking account of intra-sectoral externalities, MSC then represents marginal social costs; when we also account for the marginal environmental external costs MEC, TMSC gives the 'total marginal social costs'. Optimal road usage is then found at N^*, where net social benefit (the area between the curves MPB and TMSC) is maximised, and

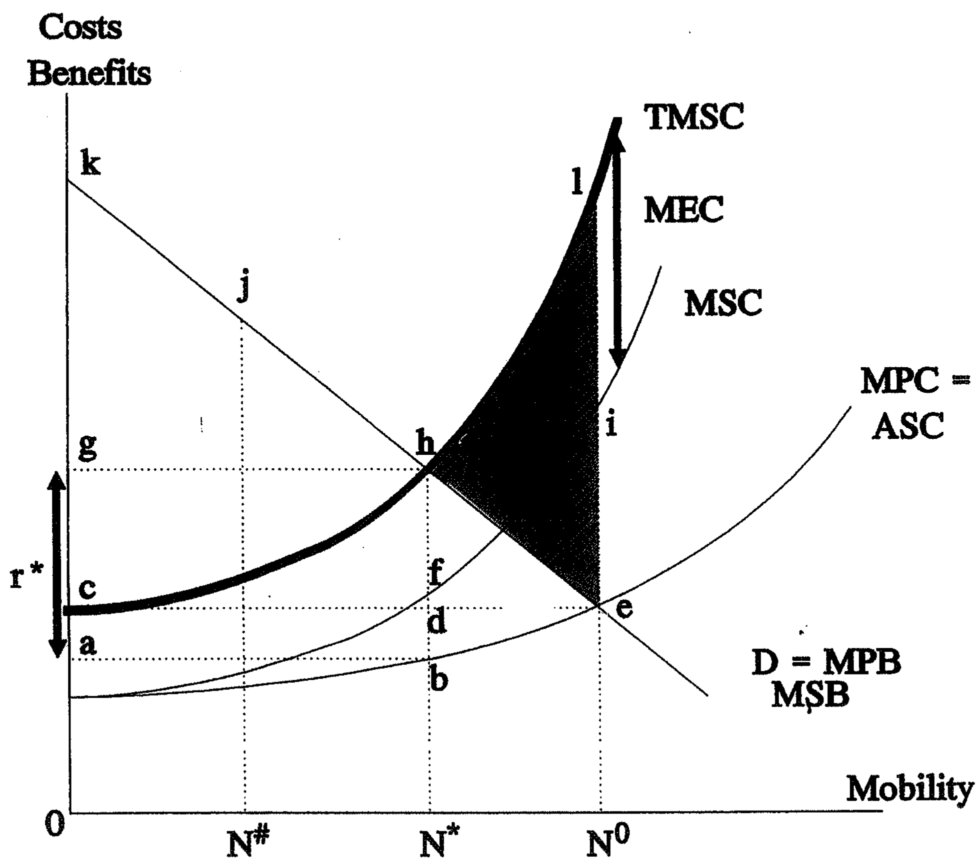

Figure 1 Welfare implications of regulation.

the shaded welfare loss, hel, is avoided. Although diagrams such as Figure 1 are usually taken to present the situation on a certain link on a certain time of day, the figure can also be seen as an abstraction for the more general road transport case.

The identification of N^* as 'optimal' is of course contingent on the welfare criterion applied. The above used principle of considering total welfare (as the sum of individuals' welfare) is based on the potential Pareto criterion, associated with Kaldor[20] and Hicks.[21] In terms of Figure 1, this criterion classifies any change in mobility as an improvement as long as the winners could compensate the losers in such a way that eventually everyone is better off (or at least nobody is worse off). The question of whether such compensations actually do take place is considered as a political matter, not relevant for the change's evaluation.[22]

This more or less commonly accepted potential Pareto criterion to a considerable extent bypasses issues of the 'social feasibility' of government policies, which is not so much dependent on the question of whether society at large benefits, but rather on the distribution of such a (net) welfare improvement – expressed, for instance, in the numbers of winners and losers, combined with the intensities of individual welfare changes. For example, in a very simple democracy, where all decisions are taken by referendum, a rule of thumb might be that at least half of the (voting) population should benefit from a certain policy (change), as otherwise it would not be accepted. Of course, most democracies do not operate in such a simple manner. Nevertheless, comparable decision- and policy mechanisms will be at hand. At the end, there will be a certain limit to the freedom of a democratically elected government (aiming at being re-elected) in the choice of their (regulatory) policies.

To illustrate this, we consider two 'textbook' instruments for achieving N^* in Figure 1: optimal physical regulation (a prohibition on mobility between N^* and N^0), and the optimal effluent fee r^*.[23] We assume for the moment that both policies succeed in achieving the optimum point N^*. Table 1 then shows the welfare effects of both policies for each group as the size of corresponding areas in Figure 1.

Table 1 The welfare effects of optimal physical regulation and optimal effluent fees

	Road users: $0 - N^*$	*Road users: $N^* - N^0$*	*Victims of the environmental externality*	*Regulator*	*Social (total)*
Optimal physical regulation	+ abdc	– beh + beif – abdc	+ filh	0	+ hel
Optimal regulatory fees	– cdhg = abdc – abhg	– beh + beif – abdc	+ filh	+ abhg	+ hel

With optimal physical regulaton, the road users generating the optimal mobility N^* will enjoy a welfare gain due to decreased congestion costs. This gain is represented by the rectangular abdc. The mobility foregone, between N^* and N^0, implies a loss of benefits equal to area N^*N^0eh and a reduction in private costs of N^*N^0eb, yielding a negative subtotal of beh. In addition, external congestion costs within this group will disappear, equal to the total reduction of external congestion costs, beif, minus the fraction abdc enjoyed by the remaining road users. Thus the total welfare effect given in the second column in Table 1 results. Finally, the reduction in the environmental externality implies a welfare gain to its 'victims' equal to filh. Summing now over the three groups, a social welfare gain equal to hel can then indeed be derived.

Optimal regulatory taxation yields identical welfare effects for the mobility foregone, for the victims of the environmental externality, and for society at large. However, the remaining road users are now worse off, as total tax renevues, abhg, (necessarily) exceed the reduction in congestion costs, abdc. These tax revenues of course accrue to the regulator, or more generally the government.

Therefore, in this stylized setting, where the two instruments are equally efficient in terms of accomplishing N^*, they are certainly not equivalent in terms of social feasibility. The road users generating optimal mobility enjoy a welfare gain with physical regulation, whereas they are worse off with regulatory fees. Since the other groups are likely to be indifferent between both policies, physical regulation will be more socially feasible than regulatory taxation. If, in the case sketched above, we assume the ratio of mobility and the number of road users to be constant along the horizontal axis, a majority of road users would even be in favour of physical measures, while all road users would be opposed to the effluent fee (an important assumption here is that internal (time) costs do not differ among road users).

It is important to stress that we implicitly assume that the tax revenues remain with the regulator, and that the various groups in society do not consider the possibility of benefitting from possible (re-)allocations of these financial means. Given the usual response to regulatory taxation ("the car as a cash-cow" or "yet another tax", instead of: "more money for beneficial public projects"...), governments really should formulate convincing policy packages if such scepticism of tax payers is to be overcome. Indeed, it is – in theory – by definition always possible to construct a redistribution of means (including the tax revenues) such that everyone is better off after optimal regulation. This, however, might involve taxation of those benefitting from the reduced environmental externality.[24] In Section 4, we will come back to the possible allocations of regulatory tax revenues.

However, apart from differing in terms of social feasibility, different regulatory instruments will usually also differ in terms of efficiency. In fact, the above assumption of both instruments achieving optimality is quite unrealistic. This particularly holds for physical regulation: it is hard to envisage a regulator applying 'optimal' physical regulation by identifying and prohibiting the socially excessive mobility between $N^*$and N^0. In reality, with physical regulation, the regulator runs the risk of also affecting mobility with relatively high economic benefits. In contrast to fees, which will naturally make the road users give up mobility between $N^*$and N^0 (where benefits fall short of the sum of the internal costs and the fee), a 'quasi-optimal' physical measure might in the worst case affect mobility between O and $N^\#$ (with $N^\#=N^0-N^*$). Regulation may then even be inferior to non-intervention, like in the case sketched where the benefits foregone, $ON^\#jk$, exceed the savings in social costs, N^*N^0lh.[25] Such adverse effects may for instance occur with a physical measure such as the 'car-free Sunday', temporarily applied in The Netherlands during the first oil crisis in 1973, and now again frequently mentioned in the public debate on transport and the environment. It is quite possible that road users would much rather give up part of the mobility generated on week-days rather than on Sundays. If the benefits associated with some mobility generated on Sundays exceed those associated with some mobility generated on other days, such a measure is clearly not optimal from an efficiency perspective. In conclusion, with physical regulation, there is no guarantee that the remaining mobility represents the highest benefits.

This may in turn negatively affect the policy's social feasibility. Although the majority of the initial road users may still benefit from some physical measure, the benefits foregone are higher than in the optimal case. While a referendum amongst road users may therefore still result in acceptance, the higher these benefits foregone, the more likely the measure is to provoke resistance and protest.

The main conclusion to this basic analysis therefore is perhaps a bit discouraging: there appears to be some inverse relation between the efficiency and the social feasibility of regulatory instruments for containing road transport externalities.

3. THE EFFICIENCY, EFFECTIVENESS AND SOCIAL FEASIBILITY OF REGULATION IN ROAD TRANSPORT

Although the model discussed in the previous section demonstrates the tension between the efficiency and the feasibility of regulation in road transport, it is clearly too simple to allow for a broad evaluation of the various possible instruments. This is in the first place partly due to the fact that the road transport sector is a complex one, also from the viewpoint of regulation. In the previous section, we focused on some complications resulting from the market's atomistic structure: there are many, in comparison to the market's size relatively small sources of externalities. These externalities, in turn, will be partly intra-sectoral and partly environmental in nature. Furthermore, the generators are by definition mobile, which gives rise to additional practical difficulties of regulation. Next, road transport externalities are often place and time dependent (congestion, noise, smog) and may in addition vary among users when depending on vehicle technology or driving style. A higher order problem is that the demand for road transport is often a derived demand. This implies that, unless sufficient adaptations take place in the underlying factors, regulation may be frustrated by inelasticity of demand.

Secondly, rather than just two archetypical instruments discussed in the previous section, there is a whole range of regulatory instruments for containing road transport externalities. In this section, we will discuss the effectiveness and efficiency (Section 3.2) and social feasibility (Section 3.3) of these instruments. First of all, however, we will consider the concepts of effectiveness, efficiency and feasibility as such in greater depth in Section 3.1.

3.1 *Effectiveness, Efficiency and Feasibility of Regulation*

From this paper's perspective, one may distinguish three main criteria for the evaluation of regulatory instruments. First, *effectiveness* refers to the extent to which an instrument is capable of accomplishing a certain goal (be it 'optimal mobility', or any other – perhaps less ambitions – target). Second, *efficiency* refers to the extent to which the accomplishment of such a goal yields the highest possible net social benefits, defined as the difference between social benefits and social costs. Hence, in the sequel, 'efficiency' does not merely refer to the binary question of 'to be or not to be Pareto efficient', but is instead used as a more continuous criterion (which is the very reason for distinguishing between the effectiveness and the efficiency of regulation). Clearly, effectiveness and efficiency are closely related: when a certain goal can be reached against relatively low social costs (efficiency), the realization of that goal (effectiveness) may generally be easier. Still, the distinction can be useful in practice, since high effectiveness does not necessarily imply high efficiency (consider, for instance, the complete prohibition of road transport), nor does the reverse necessarily hold (see the discussion on attitude policies in Section 3.2). Thirdly, *feasibility* refers to all factors influencing the possibility and probability of the introduction and application of a certain policy instrument. Table 2 provides an overview of factors underlying these three main criteria. It would take too long to discuss each of these factors separately here; in the remainder of this paper, most of them, when relevant, will be addressed. Two points, though, in particular are worth noting before proceeding.

First, space and time are important exogenous elements to Table 2. The spatial scale of a policy may be to a large extent determine its effectiveness and efficiency; for instance, for reasons of avoidability (consider fuel taxes or paking policies), as well as its feasibility (for reasons of feelings of justice). As far as the time element is concerned, the long-run effectiveness, efficiency, and feasibility of certain instruments may often be different from their short-run counterparts. In general, one would expect the former to exceed the latter, as in the long run people will have more opportunities to adapt to the regulation. In other words, the perfomance of a certain instrument in terms of the criteria mentioned in Table 2 will often depend on the question of (1) on which spatial scale the instrument is to be used, and (2) the time period considered.

Next, the distinction we make for efficiency into a narrowly and a broadly defined concept deserves some attention. The narrowly defined concept then is in line with the discussion in the previous section; for the broadly defined concept, additional factors also play a role. These factors are related to the 'cost of regulation', which may make it worthwhile not to opt for usually rather expensive 'first-best' insrtuments, but to rely on less perfect but considerably cheaper instruments instead. A topical example here is the question of whether to use an expensive electronic road pricing system for the regulation of congestion, and if so, whether to implement the system for an entire road network or to allow for some unregulated routes for reasons of cost savings.[26]

As far as the 'feasibility' of regulation is concerned, different perspectives may be taken. Apart from the social feasibility considered in the remainder of this paper, one could distinguish various other factors influencing the possibility and probability of a certain policy to be introduced. The *economic feasibility* refers to the extent to which net social benefits can be expected from a policy, and is therefore directly related to the broadly defined efficiency and effictiveness. Also the *technical feasibility* is a quite straightforward concept and needs no further comments. The *institutional feasibility* is closely connected to the level of organization of winners and losers of a policy, and the relative strength of these interest groups. When considering the regulation of environmental externalities of road usage, one could here for instance think of the question of to what extent interest groups – such as the road lobby, the automobile industry, industrial interest groups, and environmental pressure groups – succeed in affecting both public opinion and the political process. Also, the extent to which these groups compete or collaborate is an important matter here. The *legislative feasibility* refers to the extent to which a certain policy is compatible with the existing legal framework. A recent example underlining the importance of such factors is given by the fact that the experiment with a separate car-pool lane in The Netherlands was eventually terminated by the court for the reason that the whole concept of 'car-pool lanes' had no legal interpretation. Finally, the *political feasibility* includes all factors that might influence the political decision making process. Apart from the other feasibility concepts, all of which may be expected to play a role at some stage in the political process, personal perceptions of certain politicians may be crucial for the political feasibility of regulation. Also, party-political considerations (for instance, related to the interests of a party's traditional adherence, to a party's general ideology, to voting deals and swaps in parliament, and even to parliamentary presence during votes) may eventually tip the balance when it comes to actual policy formulation and decision making. Finally, especially for relatively open countries like The Netherlands, international constraints such as imposed by the EU or other international institutions (the GATT) may pose restrictions on the options open for the regulation of transport.

For the *social feasibility* (which refers to the extent to which resistance from the public at large against an instrument is (or can be) avoided), again a distinction is made into a narrowly and a broadly defined concept. The narrowly defined concept

Table 2 Possible (sub-)criteria for the evaluation of regulatory instruments for containing road transport externalities.

Effectiveness	*Efficiency*		*Feasibility*		
	Narrow efficiency	*Broad efficiency (additional factors)*	*Narrow social feasibility*	*Broad social feasibility (additional factors)*	*Other feasibility concepts*
* Avoidability * Controllability * Counter-productive side- effects * Reliability * Adaptation possibilities (to remain effective with expected growth of road mobility)	* Differentiation according to individual benefits * Differentiation according to individual external costs: vehicle, time of driving, route, trip length, driving style * Avoidability * Controllability * Unwarranted or counterproductive side-effects * Reliability	* Implementation costs * Application costs * Maintenance costs	* Numbers of winners and losers * Intensity of welfare changes, accounting for contingent direct or indirect compensation of welfare losses (including regulatory taxes)	* User-friendliness * Transparency * Allocation of contingent tax revenues * Effect on income distribution * Justice * Anonymity and privacy * Way of introduction * Adjustment possibilities for road users * Familiarity * Economic feasibility * Technical feasibility * Institutional feasibility	*Economic feasibility* * Broad efficiency and effectiveness *Technical feasibility* * Availability of technologies * Reliability *Institutional feasibility* * Level of organization (lobbies) of winners and losers * Relative strength of interest groups *Legislative feasibility* * Compatibility with legal framework *Political feasibility* * All other feasibility concepts * Personal perceptions of politicians * Party-political factors * International political constraints

is in line with the discussion in Section 2; for the broadly defined concept, additional factors also become relevant. These are for instance factors of a psychological, sociological or cultural nature. In addition, economic, technical and institutional feasibility may also have a major impact on the question of whether a certain policy will eventually be acceptable to the public at large.

It is clear from Table 2 that the different feasibility concepts are certainly not independent; they merely relect the primary perspective taken when considering issues of feasibility. In the sequel, we confine ourselves to the welfare economic perspective, and therefore to issues of effectiveness, efficiency and social feasibility. We now turn to a qualitative evaluation of the various possible forms of regulation of road transport externalities according to these three criteria.

3.2 *The Efficiency and Effectiveness of Various Regulatory Instruments*

Table 3 gives an overview of the main possible regulatory instruments for containing road transport externalities. A rough distinction is made into three main policy fields: 'direct' demand management, 'indirect' demand management and 'supply side' oriented policies. Below, it will be argued that these three main fields should be seen as complementary rather than as purely substitutable. Therefore, a meaningful comparison between instruments from different main fields in terms of efficiency and effectiveness is not really possible, since the overall performance of regulation will to a considerable extent depend on the coordination of instruments originating from these three different main fields. Furthermore, it should be noted that many instruments may actually have impacts in more than one main field, both because of substituting behaviour of road users (for instance, theoretically speaking, 'first-best policies in the first column will lead to optimal behaviourially responses in each of the three main fields), and because of the nature of the instruments as such (note, for instance, the occurence of speed and fuel policies in the first and the third main field.)

The first two columns of Table 3 contain pricing instruments, comparable to the 'fee' in the previous section. Still, these instruments certainly do differ in terms of efficiency. From an efficiency point of view, perfect marginal external cost pricing is optimal, by providing the efficient incentives to simultaneously change behaviour in each of the three mian fields.[27] The narrow efficiency therefore is optimal. Such 'first-best' policies, however, will usually be difficult or even impossible to conduct in practice, for reasons as given in the introduction to this section (for instance, the fact that marginal external costs often vary with time, place, route, vehicle, and driving style). Therefore, we will use the concept of 'first-best' more as a theoretical benchmark, than as a practical policy alternative.

The pricing instruments in the second column are therefore classified as 'second-best', since perfect fee differentiation, and hence perfect incentive provision, is not possible. Welfare losses in comparison with 'first-best' regulation are practically unavoidable. For instance, with parking fees, it is hardly possible to differentiate according to trip length or vehicle used, which renders such a policy a quite inefficient tool for containing environmental externalities. On the other hand, parking fees may be quite efficient for the regulation of urban congestion (which is strongly time and place dependent). In contrast, fuel taxes may be quite efficient for the regulation of environmental externalities (as long as not strongly time and place dependent),[28] whereas they will perform poorly for the regulation of congestion. The extent then to which the instruments in the second column approach the 'first-best' ideal, strongly depends on the extent to which the 'second-best' fee does provide the correct incentives by differentiating according to marginal exttemal costs generated.[29]

Table 3 Various regulatory instruments for containing road transport externalities

First-best policies	Direct demand management			Indirect demand management	Supply side oriented policies
	Second-best demand	*Third-best demand mamangement*	*Other direct demand policies*		
* Marginal external cost pricing, e.g. by means of: – Perfectly fluctuating electronic road pricing with perfect information provision – Perfect emission charges – Optimal combination of second-best instruments * Differentiated or simple system of tradable permits	* Imperfect electronic road pricing * Parking fees * Toll booths * Fuel taxes (effect on road transport demand) * Peak hour permits * Cordon charges * Area licenses	* Restriction of parking space * Car-free (sun)days * Minimal vehicle occupancy * Odd/even number plates on odd/even days only * Speed limits (effect on demand) * Traffic calming by means of physical characteristics of road infrastructure	* Attitude and persuasion policies, for instance, by means of public campaign	* Stimulation of alternative transport modes, such as public transport * Physical planning and spatio-economic policies * Spreading of working and shopping hours * Stimulation tele-activities	* Technology policies (vehicles) * Fuel policies: fuel quality, and effect on fuel efficiency of vehicles * Speed policies (effect on safety and emissions) * Infrastructural policies (effect on congestion and safety) * Telematics policies (efficiency of road usage)

Although 'second-best' policies may have a smaller than optimal narrow efficiency, the costs of regulation are likely to be lower than those for instruments more closely approaching 'first-best' standards (such as electronic road pricing with perfect information provision, or ingenious pollution pricing mechanisms), which will positively affect the broad efficiency. Finally, when combined well, these instruments may approach the optimal fee differentiation when considering all externalities of road usage. This is the reason for mentioning this possibility in the first column.

In addition, the effectiveness of the pricing instruments in the second column will usually fall short of 'first-best' standards. In particular, it may be possible to avoid the instruments (for instance, private parking or through-traffic in case of parking fees; or trans-boarder fuel purchases – often implying increased mobility – with fuel taxes).

Divided over the first two columns, there is a somewhat different group of pricing instruments: tradeable permits. Such systems can be envisaged in various forms (vehicle-kilometre permits, fuel permits, tradeable peak-hour smart cards for ERP systems), where the narrow efficiency will to a considerable extent depend on the degree of differentiation (area, time of driving, etc.) within the system. For the broad efficiency, it will in particular be important to keep transaction costs down.

We classify the instruments in the third column as 'third-best' because the price mechanism is not used here. These instruments are therefore instances of what was called 'physical regulation' in Section 2. Apart from a lower narrow efficiency, these instruments can be expected to be (again) less effective than 'second-best' demand management policies; due to, for instance, easier avoidability (for example, rescheduling with car-free Sundays), and unwarranted or even counterproductive side-effects (such as increased driving around in search of a parking spot after a restriction in parking space supply). However, the broad efficiency of these instruments will generally be positively affected by the absence of tolling mechanisms, making the costs of regulation generally lower. Speed policies and infrastructural measures, insofar as directed to demand reduction, are somewhat out of place in this column. With these instruments, road users pay a 'fee' in terms of time rather than money. However, since these time costs are bygone, and do not generate redistributable financial means, these instruments are filed under 'third-best' rather than 'second-best' demand management.

The fourth and fifth columns of Table 3 contain a number of regulatory instruments which are not directly comparable to those discussed on Section 2. These policies aim at accomplishing favourable (inward) shifts of the demand curve, rather than movements along this curve.

This may, for instance, be aimed at by means of direct instruments, such as attitude and persuasion policies. Without the support of other forms of direct demand management, these instruments – mentioned in the fourth column – are not likely to be very effective, as their impact is completely contingent on voluntary response.[30] Still, insofar as these instruments do succeed in reducing road transport demand, this can be expected to take place in a quite efficient way, as road users themselves can decide whether to be influenced, implying that mobility with relatively high benefits will generally remain unaffected. We have to be careful in this respect, however, since it is not clear whether the preferences before or after the persuasion should be taken as a frame of reference. A very convenient assumption to make would be that the ordering of the vehicle-kilometres along the horizontal axis of Figure 1 remains unaffected due to persuasion. Then, such policies can be classified as efficient in the narrow definition.

A rather different group of policies aiming at inward shifts of the demand curve for road transport is given by 'indirect demand management' in the fifth column, such as: the stimulation of alternative, less distortive modes of transport (cycling, public transport, etc.);[31] spatial-economic policies;[32] peak flattening through adaption of working or shopping hours; or trip suppression by stimulation of tele-activities.[33] Such instruments may be quite effective in the reduction (or spreading) of the demand for road transport, particularly since they affect factors behind this derived demand. Moreover, without using these instruments, the regulator may very well be confronted with a highly inelastic demand for road transport, which may frustrate efforts in direct demand management.[34]

Finally, the last column contains a number of 'supply side oriented' policies, aiming at shifting the external cost function downwards. Both financial and command-and-control regulations are possible, for instance, directed to vehicle technology (like the catalytic converter) or fuel quality (such as lead-free petrol). Insofar as speed policies have a favourable effect on emissions, safety, or congestion, they also belong to this category.

Infrastructural policies may be used to relieve congestion (be it at the risk of provoking latent demand, adversely affecting environmentally externalities), may be directed to issues of road safety, or could be used to give priority to certain types of road usage (for example, bus lanes, car-pool lanes or car-pool parking facilities). Finally, new telematic technologies can be used to relieve non-recurrent congestion effects.[36] These supply side oriented policies may be rather effective in reducing the externalities per vehicle-kilometre. However, such gains often evaporate because of volume effects, which give rise to the necessity of combining these instruments with demand management policies. With (perhaps) the exception of safety, this holds generally for congection as well as for environmental externalities.

This brings us to the existence of interdependencies and interactions between the various instruments mentioned. First of all, optimal 'first-best' marginal external cost pricing (in the first column) will in theory provide road users with the optimal incentives to adjust their behaviour in terms of, for instance, modal choice (fifth column), car pooling (fourth column), cleaner fuel and/or vehicles (sixth column), trip timing (fifth column), and adjustments in the locations of their various activities (fifth column). However, even when applying 'first-best' demand management policies, it is likely that government should still have to play an important role in the other two main fields. This is particularly due to the likely existence of various types of market failures and rigidities; for instance, in the supply of infrastructure, in the supply of alternative (often public) transport services, in the organisation of car-pool schemes, in research and development in fuel and vehicle technologies, or in spatial-economic development and physical planning. In general, the less efficient and effective the instruments in the first main field are used, the more important stringent policies in the other two main fields will be. Conversely, policies in these latter two main field will be more effective, the more road users – through direct demand management – receive incentives to adjust their behaviour in the preferred direction.

Ideally, one would aim at a balanced development of the three main fields distinguished in Table 3 (direct demand characteristics, factors behind that demand, and factors on the supply side). This implies that in policy formulation, one should give serious considerations to the interactions between these fields, both via market processes and because of interdependencies between the instruments as such.[37] The overall efficiency and effectiveness of the policy package will first and foremost depend on the consistency of policies from these three main fields. It is worth stressing again that instruments from different main fields can generally not be seen as pure substitutes. Only in very extreme cases, where externalities completely vanish because of a certain policy (for instance, if an entirely environmentally friendly, silent, safe and 'congestion free' car were to be developed), the necessity of such coordination would disappear.

3.3 *The Social Feasibility of Various Regulatory Instruments*

Having considered issues of efficiency and effectiveness, we now turn to a qualitative discussion of the social feasibility of the various instruments listed in Table 3.

Let us begin with the first main field: direct demand management. For the comparision of 'first-best' and 'second-best' pricing instruments, we use one of the conclusions from a related piece of research by the authors,[12] which is that the optimal 'second-best' fee is a weighted average of the optimally differentiated 'first-best' fees. Therefore, for road users generating relatively low externalities (for instance, driving relatively clean vehicles), the narrowly defined feasibility (see Section 3.1) of 'first-best' effluent taxation will exceed that of 'second-best' regulation. The opposite of course holds for road users generating relatively high externalities. This may have a positive effect on the broad social feasibility of 'first-best' taxation (in comparison with 'second-best' taxation), as it is less likely to provoke feelings of injustice. On the other hand, the broad social feasibility of an instrument such as electronic road pricing may be negatively affected for reasons of privacy and unfamiliarity. In addition, its introduction may very well be interpreted as "an additional form of taxation", which may induce stronger resistance than does an increase of existing fees (such as fuel taxes or parking charges). Hence, neither for the narrowly nor the broadly defined feasibility concept can a straightforward ranking of 'first-best' and 'second-best' effluent taxation be given.

Both of them, however, are likely to be outperformed by 'third-best' measures when it comes to the (narrow) social feasibility. Each of the three types of regulation aims at a reduction in road usage; with pricing instruments, however, this goal is pursued by making the road users pay charges (see Section 2). In this respect, a noteworthy position is taken by 'tradeable permits'. The narrow feasibility of such schemes may considerably exceed that of other pricing instruments, because the regulator can in principle distribute the permits for free. When subsequent trading of these permits takes place, there is a zero-sum game (in terms of financial flows) for the traders involved – which is obviously not the case for effluent taxation. As the regulator can decide about the initial distribution of the permits, (s)he is equipped with a potentially strong means of affecting the narrow social feasibility of regulation. The broad social feasibility of such schemes, however, will generally be negatively affected for reasons of unfamiliarity, and does also strongly depend on the costs (including the efforts) of trading.

Finally, the group of attitude and persuasion instruments may generally be expected to have a relatively high social feasibility, since the road users' contingent response to such instruments will have to be voluntary.

For the second main field, 'indirect demand management', the social feasibility of regulation will first and foremost depend on the nature of the measures taken. Clearly, restrictive policies such as prohibitions and taxes (for example, in physical planning) are less popular and socially feasible than policies conducted in terms of stimulation and subsidization (such as the supply of additional public transport or telematics networks).

Thirdly, in a qualitative sense, instruments from the third main field of supply side oriented policies often seem to provoke relatively limited resistance, especially when comparing them to direct demand management. This may partly be explained by the fact that the former do not directly threaten the 'freedom of mobility'. On the contrary, road users may very well realise that, for instance, technology based solutions can provide some guarantee for retaining road mobility. Furthermore, insofar as technology based instruments do affect the price of mobility, road users may feel that they get some 'value for money' (a cleaner car, higher quality fuel), whereas pollution taxes are usually considered as losses from the individual point of view. Finally, instruments related to the supply, maintenance and quality of infrastructure, as well as the legislation on traffic rules and speed limits, are traditionally seen as government's responsibility, and may therefore have a relatively high (broad) social feasibility.

We conclude this section by noting that – as is the case for the efficiency and effectiveness of regulation – a comparison of the social feasibility of instruments from different main policy fields cannot be made in any straightforward manner without conducting empirical and formal research, including research on the interdependencies between these instruments. However, such a comparison and trade-off between efficiency, effectiveness and the social feasibility of regulation within the first main field of direct demand management to some extent can be made – especially when focusing on the narrowly defined concepts. This is the topic of the next section.

4. TOWARDS A TRADE-OFF BETWEEN EFFICIENCY, EFFECTIVENESS AND THE SOCIAL FEASIBILITY OF REGULATION IN ROAD TRANSPORT

In this section, we will focus on the apparent necessity to come to some trade-off between the efficiency, effectiveness and feasibility of regulation in road transport. We will confine ourselves to instruments from the first main field of direct demand management. Furthermore, we will restrict ourselves to the trade-off between the narrow social feasibility on the one hand, and a combined narrow efficiency/effectiveness measure on the other – a combination which makes sense because of the close relationship between the two.

In Figure 2, the four categories of direct demand management distinguished in Table 3 are classified according to these two criteria. Their relative positions are based on the discussion in Section 3. The most favourable area in Figure 2 is of course the upper-right hand corner, where both the narrow social feasibility, and narrow efficiency and effectiveness are relatively high.

On the left-hand side of Figure 2, the regulatory fees can be found, with a relatively low narrow social feasibility. 'First-best' fees will by definition be more narrowly efficient than 'second-best' fees, and are in addition likely to be more effective. As indicated by the arrows, however, the efficiency of 'second-best' policies approaches 'first-best' standards, when the fees can be more differentiated according to individual marginal external costs. The same holds if the policy has only limited unwarranted side effects, and if it is relatively difficult to avoid.

Tha narrow feasibility of pollution taxation (be it 'first-best' or 'second-best') can be increased by 'ear-making' or 'ring-fencing' the tax revenues, and reserving them for certain goals (which should of course be as much as possible in the interest of road users).[38] In terms of Figure 1 and Table 1, this means that road users consider a certain proportion of abhg no longer as a pure loss. Three types of such 'ear-making' will briefly be discussed below.

The most straightforward possibility is more or less direct compensation of the fees charged. The problem then of course is how to do this without eroding the regulatory impact of the fees. In the most extreme case of direct and complete restitution, the fee would have no impact at all, and the market would return to the non-intervention outcome, N^0 in Figure 1. In Figure 2, this is illustrated by the south-easterly directional arrows. This problem can only be avoided if 'lump-sum' compensation were possible, meaning that the level of individual restitution does not depend on individual behaviour (in contrast of course to the individual level of the regulatory tax sum, which should depend on individual mobility behaviour).

Unfortunately, the concept of pure lump-sum payments is mainly a theoretical benchmark, with only very limited practical relevance. The second type of 'earmarking', however, being the variabilization of road taxes, can be considered as a (not inexhaustible) form of lump-sum compensations. Such simultaneous introduction of pollution taxes

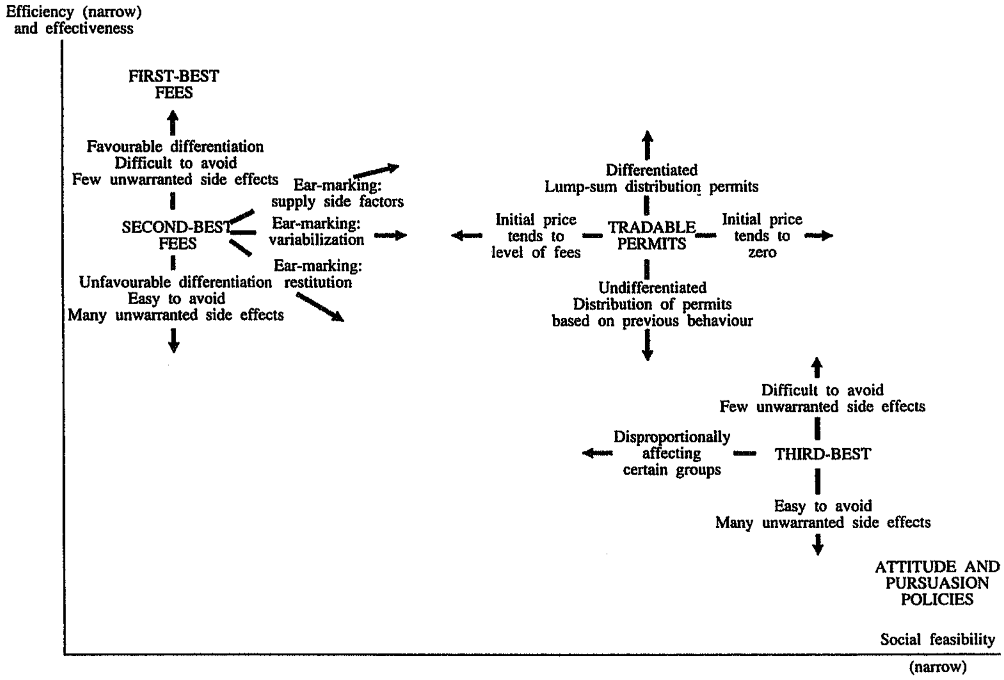

Figure 2 The trade-off between efficiency, effectiveness and social feasibility of direct demand management for the regulation of road transport.

and lowering of fixed taxes can offer an important possibility of avoiding resistance against the introduction of regulatory taxation.

The third form of 'ear-marking' is often the theoretically most elegant: the usage of regulatory tax revenues for covering the costs of measures from the third main field of supply oriented policies. For instance, for congestion, it can be shown that – under certain assumptions – the revenues of optimal congestion tolls are just sufficient to cover the cost of the optimal supply of road infrastructure.[39] A comparable analysis can be made for the revenues of environmental taxes versus the costs of investments in cleaner public technologies, such as noise barriers or investments in research and development. The proverbial knife then cuts from three sides: the regulatory impact of the fee, the increased feasibility of regulation, and the positive impacts of supply side oriented policies. In Figure 2, this is depicted by the north-easterly pointing arrows.

The horizontal position (that is, the narrow feasibility) of tradeable permits strongly depends on the question of how these permits are to be issued. If they are sold, the narrow feasibility will decline as the initial price approaches more closely the level of corresponding regulatory fees. When these two become equal, no trade will take place, as only those road users that were prepared to pay a fee would purchase a permit anyway. The two instruments are then equivalent: the narrow feasibility will strongly depend on the allocation of the tax revenues, whereas the efficiency and effectiveness will depend on the degree of diferentiation with the system. In the other extreme, where the permits' initial price tends to zero, the narrow feasibility will of course increase. However, the efficiency will then more and more depend on the question of how the permits are to be distributed. If the initial distribution depends on revealed mobility behaviour, for instance, in some previous period, then a situation comparable to direct restitution of tax revenues might arise. If road users know that a large share of the money spent on permits will be refunded in the next period in the form of free permits, strategic behaviour can easily arise, which (although it cannot affect the effectiveness) may have a negative impact on efficiency.[40] Also in this case, a lump-sum distribution of permits is preferable from an efficiency point of view.

As indicated in the figure, tradeable permits – when distributed for free – may have a narrow feasibility comparable to that of 'third-best' regulation. Moreover, this latter may even be exceeded if certain groups are disproportionally affected by such 'third-best' regulation, as outlined in Section 2 and indicated by the arrow pointing leftwards. Tradeable permits in these cases have the advantage of allowing such groups to retain their mobility by purchasing permits from those who derive less benefits from mobility. As such trade takes place voluntarily; it can only positively affect the social feasibility in comparison to 'third-best' regulation where such trade is not possible. As depicted by the vertical arrows, the efficiency and effectiveness of 'third-best' policies will depend particularly on the avoidability of and the extent to which unwarranted side effects occur.

The position of the final block of attitude and persuasion policies follows directly from the discussion in the previous section. Although quite feasible, and not particularly inefficient, such instruments will most likely turn out not to be very effective.

It is perhaps important to stress here that we have confined ourselves to the narrowly defined concepts of efficiency and feasibility. Broadening the scope to the broadly defined concepts may certainly affect the relative positions of the various instruments in Figure 2. The broad efficiency of 'first-'best' charging will often be negatively affected by the costs of differentiated toll collection. The same can be said for tradeable permits when considering transaction costs. Also the broad feasibility of these instruments may be negatively affected for reasons of unfamiliarity; whereas 'second-best' instruments will often benefit from already being used to some extent. However, in the qualitative framework presented above, it is not possible to discuss these broader concepts without speculation – which is the reason for restricting the analysis to the narrowly defined concepts.

5. CONCLUSION

In this paper we have discussed the tension between the theory and practice of the regulation of road transport by linking this to the trade-off between efficiency, effectiveness and the social feasibility of regulation. As there is a certain limit to the freedom of a democratic government in formulating their regulatory policies, this is a highly important (but yet hardly explored) research area. Unfortunately, there appears to be some inverse relation between the efficiency and effectiveness of regulation on the one hand, and its social feasibility on the other. However, the regulator is not merely restricted to making a negative trade-off between efficiency, effectiveness and the social feasibility of regulation. From a more positive perspective, there appears to be room for strategies aiming at increasing the efficiency and effectiveness of regulation, given a certain social feasibility – and vice versa.

Three important trajectories in this respect may be (1) the usage of tradeable permits, (2) the formulation of well balanced mixes of 'second-best' and 'third-best' policies, approaching 'first-best' fee differentiation as much as possible, and (3) 'earmarking' of regulatory tax revenues. In the latter case, the aim is to turn potential Pareto improvements as much as possible into strict Pareto improvements, without eroding the efficiency and effectiveness of regulation. As long as regulation leads to net social welfare gains, a strict Pareto improvement is theoretically possible. In practice, however, this requires quite some creativity in finding ways of redistribution that are as much as possible lump-sum in nature. Usage of tax revenues for financing supply side oriented policies and variabilization of road taxes can play an important role in this respect. From a broader perspective, the social feasibility of regulation may be increased by compensations in other taxes, as well as by making the public aware of the benefits (and perhaps necessity) of environmental improvements.

Furthermore, the coordination of policies is an important strategic weapon. This holds both for the combination of 'second-best' policies, and for the coordination between 'direct' demand management, 'indirect' demand management and 'supply side oriented' policies. This latter issue is of particular importance for the overall efficiency and effectiveness of the whole policy package, and hence also for its social feasibility (via the extent to which instruments will have to be used disproportionally stringent because of the lack of policies in the other fields).

Finally, from a long run perspective, it is important that the choice of a policy package is such that it remains capable of meeting the criteria of efficiency, effectiveness and social feasibility, also in the light of the expected continuing growth in the demand for transport. Furthermore, more stringent policies are preferably to be introduced as soon as possible – not only because of the current magnitude of road transport externalities, but especially as current mobility patterns provoke and facilitate shifts in the factors behind the demand of transport (for instance, in terms of spatial organisation) that will undoubtedly negatively affect the effectiveness and social feasibility of future regulation.

Acknowledgement

The authors would like to thank Emile Quinet and Jeroen van den Bergh for valuable comments on an earlier version of this paper.

References and notes

1. S. F. Borins, "Electronic Road Pricing: An Idea Whose Time May Never Come." *Transportation Research* **22A** (1), 37–44 (1988).
2. A. W. Evans, "Road Congestion Pricing: When Is It a Good Policy?" *Journal of Transport Economics and Policy* **26**, 213–243 (1992).
3. G. Giuliano, "An Assessment of the Political Acceptability of Congestion Pricing." *Transportation* **19** (4), 335–358 (1992).
4. C. Lave, "The Demand Curve under Road Pricing and the Problem of Political Feasibility." *Transportation Research* **28A** (2), 83–91 (1994).
5. E. T. Verhoef, "The Demand Curve under Road Pricing and the Problem of Political Feasibility: A Comment." *Transportation Research* **29A** (6), 459–464 (1995).
6. P. B. Goodwin, "The Rule of Three: A Possible Solution to the Political Problem of Competing Objectives for Road Pricing." *Traffic Engineering and Control* **30** (10), 495–497 (1989).
7. P. Jones, "Gaining Public Support for Road Pricing through a Package Approach." *Traffic Engineering and Control* **32** (4), 194–196 (1991).
8. A. D. May, "Road Pricing: An International Perspective." *Transportation* **19** (4), 313–333 (1992).
9. R. W. Poole (Jr.), "Introducing Congestion Pricing on a New Toll Road." *Transportation* **19** (4), 383–396 (1992).
10. K. A. Small, "Using the Revenues from Congestion Pricing." *Transportation* **19** (4), 359–81 (1992).
11. K. J. Button, "*Alternatives to Road Pricing*." Paper presented to the OECD/ECMT/NFP/GVF Conference on *The Use of Economic Instruments in Urban Travel Management*, Basel (1992).
12. E. T. Verhoef, P. Nijkamp and P. Rietveld, *Second-best Regulation of Road Transport Externalities. Journal of Transport Economics and Policy* **29**, 147–167 (1995).
13. H. Mohring, "The Role of Fuel Taxes in Controlling Congestion." In: *Transport Policy, Management and Technology Towards 2001*. Proceedings of the Fifth World Conference on Transport Research 1: 'The Role of Public Sector in Transport', Yokohama, 243–257 (1989).
14. R. Arnott, A. de Palma and R. Lindsey, "A Temporal and Spatial Equilibrium Analysis of Commuter Parking." *Journal of Public Economcis* **45**, 302–335 (1991).
15. A. Glazer and E. Niskanen, "Parking Fees and Congestion." *Regional Science and Urban Economics* **22**, 123–132 (1992).

16. E. T. Verhoef, P. Nijkamp and P. Rietveld, "The Economics of Regulatory Parking Policies." *Transportation Research A*. **29A** (2), 141–156 (1995) Forthcoming in Journal of Urban Economics.
17. E. T. Verhoef, P. Nijkamp and P. Rietveld, *Second-best Congestion Pricing: The Case of an Untolled Alternative*. TRACE discussion paper TI 94-129, Tinbergen Institute, Amsterdam-Rotterdam (1994).
18. E. T. Verhoef, "External Effects and Social Costs of Road Transport." *Transportation Research* **28A** (4), 273–287 (1994).
19. Significant external benefits of road transport are not likely to exist; the benefits are usually either purely internal or pecuniary in nature (see also [18]). Hence, MPB and MSB are assumed to be identical in Figure 1.
20. N. Kaldor, "Welfare Propositions of Economics and Interpersonal Comparisons of Utility." *Economic Journal* **49,** 549–552 (1939).
21. J. R. Hicks, "The Foundation of Welfare Economics." *Economic Journal* **49**, 696–712 (1939).
22. The strict Pareto criteria, where a change is considered an improvement only if also without compensation nobody is worse off (and at least one actor is better off), is not a very useful criterion in practice, because such situations hardly ever occur.
23. See, for a comparable but more extensive analysis for the regulation of environmental externalities in general, E. T. Verhoef, "Efficiency and Equity in Externalities: A Partial Equilibrium Analysis." *Environment and Planning* **26A**, 361–382 (1994).
24. Insofar as the environmental benefits are to be enjoyed by future generations, there will of course be a practical problem. To a limited extent however, this may justify the existence of public debts (in the interpretation of postponed taxation).
25. Furthermore, with physical regulation it is conceivable that, due to its favourable effect on user costs, latent demand to the fight of N^0 (insofar as it is not affected by the prohibition) will be provoked.
26. See, for a formal analysis, Verhoef, Nijkamp and Rietveld (1994) [17].
27. The reason then for filing the option of 'first best' policies under 'direct demand management' is that, with an elastic demand for vehicle miles, the first behavioural adaptations may be expected to be made in terms of reductions in mobility. Still, our reluctance to classify 'first best' policies as merely 'direct demand management' is reflected by the physical lay-out of Table 3; in particular by printing 'First-best policies' in bold face and by shifting it upwards in comparision to the headings of the other sub-categories of 'direct demand management'.
28. See also Verhoef, Nijkamp and Rietveld (1995) [16].
29. For a formal analysis, we refer to Verhoef, Nijkamp and Rietveld (1995) [12].
30. See, for a psychological analysis of voluntary car mobility reduction, G. Tertoolen, *Free to Move...?! A Field Experiment on Attempts to Influence Private Car Use and the Psychological Resistance It Evokes* (in Dutch). Ph.D. Dissertation, University of Utrecht (1994).
31. S. Rienstra, J. Vleugel and P. Nijkamp, *Options for Sustainable Passenger Transport; An Assessment of Policy Choices*. Draft, Free University, Amsterdam (1994).
32. E. T. Verhoef, J. C. J. M. van den Bergh and K. J. Button, "Transport, Spatial Economy and the Global Environment" Environment and Planning A (forthcoming).
33. P. M. Jones and I. Salomon, "Technological and Social Developments and their Implications for In-Home/Out-of-Home Interactions." In: Nijkamp, P. (ed.), *Europe on the Move: Recent Developments in European Communications and Transport Activity Research*. Avebury, Aldershot, 95–113 (1993).
34. See Verhoef, Van den Bergh and Button (forthcoming) [32].
35. E. T. Verhoef, R. H. M. Emmerink, P. Nijkamp and P. Rietveld, *Information Provision, Flat- and Fine Tolling and the Efficiency of Road Usage*. TRACE discussion paper TI 94–157, Tinbergen Institute, Amsterdam-Rotterdam (1994).
36. H. Geerlings, "The Meaning of Technology Policy in the Transport Sector and the Role of the Government; New Opportunities for New Policies" (in Dutch). In: *Colloquium Vervoersplanologisch Speurwerk 1994*. C.V.S., Delft, 53–72 (1994).
37. See Verhoef, Van den Bergh and Button (forthcoming) [32] for a formal analysis.
38. If the allocation of these revenues is not directly in the interest of road users, the feasibility may even be negatively affected. This recently happened in the Dutch province of Drenthe, where the road lobby accused the regional authorities of 'legal theft' after the news had spread that the extra revenues of increased road taxes would be used for some subsidy scheme for thatched roofs.
39. See Verhoef, Van den Bergh and Button (forthcoming) [32].
40. Only if capital markets operate perfectly, and everybody has the same perception of the expected price in the next period, need this not be a problem.